

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

L Number	Hits	Search Text	DB	Time stamp
1	3	(("6,408,290") or ("6,345,265") or ("6,556,958")).PN.	USPAT; US-PGPUB	2004/08/21 22:31
2	12	Bayesian adj network and database and expectation adj maximization	USPAT; US-PGPUB	2004/08/21 22:32


[Membership](#) [Publications/Services](#) [Standards](#) [Conferences](#) [Careers/Jobs](#)

 Welcome
 United States Patent and Trademark Office

 IEEE Xplore®
 1 Million Documents
 1 Million Users

[Help](#) [FAQ](#) [Terms](#) [IEEE Peer Review](#)
Quick Links
[» Search Results](#)

Welcome to IEEE Xplore®

- Home
- What Can I Access?
- Log-out

Tables of Contents

- Journals & Magazines
- Conference Proceedings
- Standards

Search

- By Author
- Basic
- Advanced

Member Services

- Join IEEE
- Establish IEEE Web Account
- Access the IEEE Member Digital Library

- Access the IEEE Enterprise File Cabinet

Print Format

[Home](#) | [Log-out](#) | [Journals](#) | [Conference Proceedings](#) | [Standards](#) | [Search by Author](#) | [Basic Search](#) | [Advanced Search](#) | [Join IEEE](#) | [Web Account](#) | [New this week](#) | [Linking Information](#) | [Your Feedback](#) | [Technical Support](#) | [Email Alerting](#) | [No Robots Please](#) | [Release Notes](#) | [IEEE Online Publications](#) | [Help](#) | [FAQ](#) | [Terms](#) | [Back to...](#)

Copyright © 2004 IEEE — All rights reserved


[Membership](#) [Publications/Services](#) [Standards](#) [Conferences](#) [Careers/Jobs](#)

 Welcome
 United States Patent and Trademark Office

 IEEE Xplore®
 1 Million Documents
 1 Million Users

[Help](#) [FAQ](#) [Terms](#) [IEEE Peer Review](#)
Quick Links
[» Search Results](#)

Welcome to IEEE Xplore®

- Home
- What Can I Access?
- Log-out

[Tables of Contents](#)

- Journals & Magazines
- Conference Proceedings
- Standards

[Search](#)

- By Author
- Basic
- Advanced

[Member Services](#)

- Join IEEE
- Establish IEEE Web Account
- Access the IEEE Member Digital Library

- Access the IEEE Enterprise File Cabinet

[Print Format](#)
[Home](#) | [Log-out](#) | [Journals](#) | [Conference Proceedings](#) | [Standards](#) | [Search by Author](#) | [Basic Search](#) | [Advanced Search](#) | [Join IEEE](#) | [Web Account](#) | [New this week](#) | [Linking Information](#) | [Your Feedback](#) | [Technical Support](#) | [Email Alerting](#) | [No Robots Please](#) | [Release Notes](#) | [IEEE Online Publications](#) | [Help](#) | [FAQ](#) | [Terms](#) | [Back to.....](#)

Copyright © 2004 IEEE — All rights reserved


[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)
[Search: The ACM Digital Library](#) [The Guide](#)

US Patent & Trademark Office

"Bayesian network" and database and "expectation maximizati



THE ACM DIGITAL LIBRARY

[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Terms used [Bayesian network](#) and [database](#) and [expectation maximization](#)

Found 2,561 of 141,345

Sort results by

Save results to a Binder

[Try an Advanced Search](#)

Display results

Search Tips

[Try this search in The ACM Guide](#)

Open results in a new window

Results 1 - 20 of 200

Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale



1 [Item-based top-N recommendation algorithms](#)

Mukund Deshpande, George Karypis

January 2004 **ACM Transactions on Information Systems (TOIS)**, Volume 22 Issue 1Full text available: [pdf\(240.61 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

The explosive growth of the world-wide-web and the emergence of e-commerce has led to the development of *recommender systems*--a personalized information filtering technology used to identify a set of items that will be of interest to a certain user. User-based collaborative filtering is the most successful technology for building recommender systems to date and is extensively used in many commercial recommender systems. Unfortunately, the computational complexity of these methods grows $\text{I} \dots$

Keywords: e-commerce, predicting user behavior, world wide web



2 [Probabilistic query models for transaction data](#)

Dmitry Pavlov, Padhraic Smyth

August 2001 **Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining**Full text available: [pdf\(958.33 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

We investigate the application of Bayesian networks, Markov random fields, and mixture models to the problem of query answering for transaction data sets. We formulate two versions of the querying problem: the query selectivity estimation (i.e., finding exact counts for tuples in a data set) and the query generalization problem (i.e., computing the probability that a tuple will occur in new data). We show that frequent itemsets are useful for reducing the original data to a compressed representa ...

3 [Special issue on the fusion of domain knowledge with data for decision support: Fusion of domain knowledge with data for structural learning in object oriented domains](#)

Helge Langseth, Thomas D. Nielsen

December 2003 **The Journal of Machine Learning Research**, Volume 4Full text available: [pdf\(227.18 KB\)](#) Additional Information: [full citation](#), [abstract](#), [index terms](#)

When constructing a Bayesian network, it can be advantageous to employ structural learning algorithms to combine knowledge captured in databases with prior information

provided by domain experts. Unfortunately, conventional learning algorithms do not easily incorporate prior information, if this information is too vague to be encoded as properties that are local to families of variables. For instance, conventional algorithms do not exploit prior information about repetitive structures, which are ...

4 Posters: Combining speech and haptics for intuitive and efficient navigation through image databases 

Thomas Käster, Michael Pfeiffer, Christian Bauckhage

November 2003 **Proceedings of the 5th international conference on Multimodal interfaces**

Full text available:  pdf(239.65 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Given the size of todays professional image databases, the standard approach to object- or theme-related image retrieval is to interactively navigate through the content. But as most users of such databases are designers or artists who do not have a technical background, navigation interfaces must be intuitive to use and easy to learn. This paper reports on efforts towards this goal. We present a system for intuitive image retrieval that features different modalities for interaction. Apart f ...

Keywords: content-based image retrieval, fusion of haptics, multimodal interface evaluation, speech, vision processing

5 Multi Relational Data Mining (MRDM): Probabilistic logic learning 

Luc De Raedt, Kristian Kersting

July 2003 **ACM SIGKDD Explorations Newsletter**, Volume 5 Issue 1

Full text available:  pdf(1.98 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

The past few years have witnessed an significant interest in probabilistic logic learning, i.e. in research lying at the intersection of probabilistic reasoning, logical representations, and machine learning. A rich variety of different formalisms and learning techniques have been developed. This paper provides an introductory survey and overview of the state-of-the-art in probabilistic logic learning through the identification of a number of important probabilistic, logical and learning concept ...

Keywords: data mining, inductive logic programming, machine learning, multi-relational data mining, probabilistic reasoning, uncertainty

6 Video retrieval: Semi-supervised learning for facial expression recognition 

Ira Cohen, Nicu Sebe, Fabio G. Cozman, Thomas S. Huang

November 2003 **Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval**

Full text available:  pdf(341.70 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Automatic classification by machines is one of the basic tasks required in any pattern recognition and human computer interaction applications. In this paper, we discuss training probabilistic classifiers with labeled and unlabeled data. We provide an analysis which shows under what conditions unlabeled data can be used in learning to improve classification performance. We discuss the implications of this analysis to a specific type of probabilistic classifiers, Bayesian networks, and propose a ...

Keywords: Bayesian networks, facial expression recognition, semi-supervised learning

7

Modeling dependencies in protein-DNA binding sites 

Yoseph Barash, Gal Elidan, Nir Friedman, Tommy Kaplan

April 2003 **Proceedings of the seventh annual international conference on Computational molecular biology**

Full text available:  pdf(411.94 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The availability of whole genome sequences and high-throughput genomic assays opens the door for *in silico* analysis of transcription regulation. This includes methods for discovering and characterizing the binding sites of DNA-binding proteins, such as transcription factors. A common representation of transcription factor binding sites is a *position specific score matrix* (PSSM). This representation makes the strong assumption that binding site positions are independent of each other ...

Keywords: DNA sequence motifs, bayesian networks, factors binding sites, transcription

8 Multi Relational Data Mining (MRDM): Biological applications of multi-relational data mining

David Page, Mark Craven

July 2003 **ACM SIGKDD Explorations Newsletter**, Volume 5 Issue 1

Full text available:  pdf(1.12 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Biological databases contain a wide variety of data types, often with rich relational structure. Consequently multi-relational data mining techniques frequently are applied to biological data. This paper presents several applications of multi-relational data mining to biological data, taking care to cover a broad range of multi-relational data mining techniques.

9 Survey articles: Data mining for hypertext: a tutorial survey

Soumen Chakrabarti

January 2000 **ACM SIGKDD Explorations Newsletter**, Volume 1 Issue 2

Full text available:  pdf(1.19 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

With over 800 million pages covering most areas of human endeavor, the World-wide Web is a fertile ground for data mining research to make a difference to the effectiveness of information search. Today, Web surfers access the Web through two dominant interfaces: clicking on hyperlinks and searching via keyword queries. This process is often tentative and unsatisfactory. Better support is needed for expressing one's information need and dealing with a search result in more structured ways than av ...

10 Context-specific Bayesian clustering for gene expression data

Yoseph Barash, Nir Friedman

April 2001 **Proceedings of the fifth annual international conference on Computational biology**

Full text available:  pdf(233.32 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The recent growth in genomic data and measurement of genome-wide expression patterns allows to examine gene regulation by transcription factors using computational tools. In this work, we present a class of mathematical models that help in understanding the connections between transcription factors and functional classes of genes based on genetic and genomic data. These models represent the joint distribution of transcription factor binding sites and of expression levels of a gene in a single ...

11 Towards automated synthesis of data mining programs

Wray Buntine, Bernd Fischer, Thomas Pressburger

August 1999 **Proceedings of the fifth ACM SIGKDD international conference on**

Knowledge discovery and data mining

Full text available:  pdf(637.67 KB) Additional Information: [full citation](#), [references](#), [index terms](#)

12 Protocols: Similarity measure and instance selection for collaborative filtering 

Chun Zeng, Chun-Xiao Xing, Li-Zhu Zhou

May 2003 **Proceedings of the twelfth international conference on World Wide Web**

Full text available:  pdf(306.48 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Collaborative filtering has been very successful in both research and applications such as information filtering and E-commerce. The k-Nearest Neighbor (KNN) method is a popular way for its realization. Its key technique is to find k nearest neighbors for a given user to predict his interests. However, this method suffers from two fundamental problems: sparsity and scalability. In this paper, we present our solutions for these two problems. We adopt two techniques: a matrix conversion method for ...

Keywords: collaborative filtering, instance selection, similarity measure

13 From promoter sequence to expression: a probabilistic framework 

Eran Segal, Yoseph Barash, Itamar Simon, Nir Friedman, Daphne Koller

April 2002 **Proceedings of the sixth annual international conference on Computational biology**

Full text available:  pdf(3.22 MB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

We present a probabilistic framework that models the process by which transcriptional binding explains the mRNA expression of different genes. Our joint probabilistic model unifies the two key components of this process: the prediction of gene regulation events from sequence motifs in the gene's promoter region, and the prediction of mRNA expression from combinations of gene regulation events in different settings. Our approach has several advantages. By learning promoter sequence motifs that ar ...

14 A discriminative model for identifying spatial cis-regulatory modules 

Eran Segal, Roded Sharan

March 2004 **Proceedings of the eighth annual international conference on Computational molecular biology**

Full text available:  pdf(280.23 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Transcriptional regulation is mediated by the coordinated binding of transcription factors to the upstream region of genes. In higher eukaryotes, the binding sites of cooperating transcription factors are organized into short sequence units, called cis-regulatory modules. In this paper we propose a method for identifying modules of transcription factor binding sites in a set of co-regulated genes, using only the raw sequence data as input. Our method is based on a novel probabilistic model that ...

Keywords: cis-regulatory module, probabilistic model, transcriptional regulation

15 Motion texture: a two-level statistical model for character motion synthesis 

Yan Li, Tianshu Wang, Heung-Yeung Shum

July 2002 **ACM Transactions on Graphics (TOG) , Proceedings of the 29th annual conference on Computer graphics and interactive techniques**, Volume 21 Issue 3

Full text available:  pdf(5.06 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

In this paper, we describe a novel technique, called motion texture, for synthesizing

complex human-figure motion (e.g., dancing) that is statistically similar to the original motion captured data. We define motion texture as a set of motion textons and their distribution, which characterize the stochastic and dynamic nature of the captured motion. Specifically, a motion texton is modeled by a linear dynamic system (LDS) while the texton distribution is represented by a transition matrix indicat ...

Keywords: linear dynamic systems, motion editing, motion synthesis, motion texture, texture synthesis

16 Hypertext data mining (tutorial AM-1) 

Soumen Chakrabarti

August 2000 **Tutorial notes of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  pdf(1.08 MB) Additional Information: [full citation](#), [index terms](#)

17 Web search 1: Searching web databases by structuring keyword-based queries 

Pável Calado, Altigran S. da Silva, Rodrigo C. Vieira, Alberto H. F. Laender, Berthier A. Ribeiro-Neto

November 2002 **Proceedings of the eleventh international conference on Information and knowledge management**

Full text available:  pdf(204.22 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

On-line information services have become widespread in the Web nowadays. However, Web users are non-specialized and have a great variety of interests. Thus, interfaces for Web databases must be simple and uniform. In this paper we present an approach, based on Bayesian networks, for querying Web databases using keywords only. According to this approach, the user inputs a query through a simple search-box interface. From the input query, one or more plausible structured queries are derived and su ...

Keywords: query structuring, structured queries, web databases

18 Industry track papers: On the potential of domain literature for clustering and Bayesian network learning 

Peter Antal, Patrick Glenisson, Geert Fannes

July 2002 **Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  pdf(1.10 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Thanks to its increasing availability, electronic literature can now be a major source of information when developing complex statistical models where data is scarce or contains much noise. This raises the question of how to integrate information from domain literature with statistical data. Because quantifying similarities or dependencies between variables is a basic building block in knowledge discovery, we consider here the following question. Which vector representations of text and which st ...

Keywords: Bayesian networks, clustering, data mining, text mining

19 Poster papers: Mining complex models from arbitrarily large databases in constant time 

Geoff Hulten, Pedro Domingos

July 2002 **Proceedings of the eighth ACM SIGKDD international conference on**

Knowledge discovery and data mining

Full text available: [!\[\]\(6e934896f25e6ce1b0dbb50c23abc197_img.jpg\) pdf\(853.58 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

In this paper we propose a scaling-up method that is applicable to essentially any induction algorithm based on discrete search. The result of applying the method to an algorithm is that its running time becomes independent of the size of the database, while the decisions made are essentially identical to those that would be made given infinite data. The method works within pre-specified memory limits and, as long as the data is iid, only requires accessing it sequentially. It gives anytime resu ...

Keywords: Bayesian networks, Hoeffding bounds, discrete search, scalable learning algorithms, subsampling

20 [Classification and browsing: Structuring keyword-based queries for web databases](#) 
Rodrigo C. Vieira, Pavel Calado, Altigran S. da Silva, Alberto H. F. Laender, Berthier A. Ribeiro-Neto

July 2002 **Proceedings of the second ACM/IEEE-CS joint conference on Digital libraries**

Full text available: [!\[\]\(ceb7cef9f9d693d102dfe501130037c6_img.jpg\) pdf\(116.95 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

This paper describes a framework, based on Bayesian belief networks, for querying Web databases using keywords only. According to this framework, the user inputs a query through a simple search-box. From the input query, one or more plausible structured queries are derived and submitted to Web databases. The results are then retrieved and presented to the user as ranked answers. To evaluate our framework, an experiment using 38 example queries was carried out. We found out that 97% of the time, ...

Keywords: bayesian belief networks, web databases, web query

Results 1 - 20 of 200

Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2004 ACM, Inc.
[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads: [!\[\]\(4e9db7091c22bfa9fd8343485308f15c_img.jpg\) Adobe Acrobat](#) [!\[\]\(9db39d841202e58e29e81cbbbe04a6ea_img.jpg\) QuickTime](#) [!\[\]\(e24ec6da88877f218cdf0974809d2b4c_img.jpg\) Windows Media Player](#) [!\[\]\(80c73450fe7ae328e9d894e77eaf34e7_img.jpg\) Real Player](#)



US Patent & Trademark Office

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)
 The ACM Digital Library The Guide

 "Bayesian network" and database and "expectation maximizati


THE ACM DIGITAL LIBRARY

[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Terms used Bayesian network and database and expectation maximization

Found 2,561 of 141,345

Sort results by

 Save results to a Binder

 Try an Advanced Search

Display results

 Search Tips

 Try this search in The ACM Guide

 Open results in a new window
Results 21 - 40 of 200 Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale

21 KDD-99 conference reports: Profiling your customers using Bayesian networks

Paola Sebastiani, Marco Ramoni, Alexander Crea

January 2000 **ACM SIGKDD Explorations Newsletter**, Volume 1 Issue 2Full text available: [pdf\(1.22 MB\)](#) Additional Information: [full citation](#), [abstract](#)

This report describes a complete Knowledge Discovery session using Bayesware Discoverer, a program for the induction of Bayesian networks from incomplete data. We build two causal models to help an American Charitable Organization understand the characteristics of respondents to direct mail fund raising campaigns. The first model is a Bayesian network induced from the database of 96,376 Lapsed donors to the June '97 renewal mailing. The network describes the dependency of the probability of resp ...

Keywords: Bayesian networks, customer profiling, missing data

22 Research sessions: text and DB: When one sample is not enough: improving text

database selection using shrinkage

Panagiotis G. Ipeirotis, Luis Gravano

June 2004 **Proceedings of the 2004 ACM SIGMOD international conference on Management of data**Full text available: [pdf\(391.26 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)

Database selection is an important step when searching over large numbers of distributed text databases. The database selection task relies on statistical summaries of the database contents, which are not typically exported by databases. Previous research has developed algorithms for constructing an approximate content summary of a text database from a small document sample extracted via querying. Unfortunately, Zipf's law practically guarantees that content summaries built this way for any rela ...

23 Automatically structured and translated queries: The effectiveness of automatically structured queries in digital libraries

Marcos André Gonçalves, Edward A. Fox, Aaron Krowne, Pável Calado, Alberto H. F. Laender, Altigran S. da Silva, Berthier Ribeiro-Neto

June 2004 **Proceedings of the 2004 joint ACM/IEEE conference on Digital libraries**Full text available: [pdf\(295.40 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Structured or fielded metadata is the basis for many digital library services, including

searching and browsing. Yet, little is known about the impact of using structure on the effectiveness of such services. In this paper, we investigate a key research question: do structured queries improve effectiveness in DL searching? To answer this question, we empirically compared the use of unstructured queries to the use of structured queries. We then tested the capability of a simple Bayesian network s ...

Keywords: bayesian networks, digital libraries, structured queries

24 Special issue on the fusion of domain knowledge with data for decision support:

Combining knowledge from different sources in causal probabilistic models

Marek J. Druzdzel, Francisco J. Diez

December 2003 **The Journal of Machine Learning Research**, Volume 4

Full text available:  pdf(140.32 KB) Additional Information: [full citation](#), [abstract](#), [index terms](#)

Building probabilistic and decision-theoretic models requires a considerable knowledge engineering effort in which the most daunting task is obtaining the numerical parameters. Authors of Bayesian networks usually combine various sources of information, such as textbooks, statistical reports, databases, and expert judgement. In this paper, we demonstrate the risks of such a combination, even when this knowledge encompasses such seemingly population-independent characteristics as sensitivity and ...

25 NSF workshop on industrial/academic cooperation in database systems

Mike Carey, Len Seligman

March 1999 **ACM SIGMOD Record**, Volume 28 Issue 1

Full text available:  pdf(1.96 MB) Additional Information: [full citation](#), [index terms](#)

26 Exploration mining in diabetic patients databases: findings and conclusions

Wynne Hsu, Mong Li Lee, Bing Liu, Tok Wang Ling

August 2000 **Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  pdf(136.53 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

27 Database session 1: querying high-dimensional data: Approximate searches: k-neighbors + precision

Sid-Ahmed Berrani, Laurent Amsaleg, Patrick Gros

November 2003 **Proceedings of the twelfth international conference on Information and knowledge management**

Full text available:  pdf(154.57 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

It is known that all multi-dimensional index structures fail to accelerate content-based similarity searches when the feature vectors describing images are high-dimensional. It is possible to circumvent this problem by relying on approximate search-schemes trading-off result quality for reduced query execution time. Most approximate schemes, however, provide none or only complex control on the precision of the searches, especially when retrieving the *k* nearest neighbors (NNs) of query poi ...

Keywords: approximate nearest-neighbor searches, multimedia databases, similarity searches

28 Selectivity estimation using probabilistic models

Lise Getoor, Benjamin Taskar, Daphne Koller

May 2001 **ACM SIGMOD Record , Proceedings of the 2001 ACM SIGMOD international conference on Management of data**, Volume 30 Issue 2Full text available:  pdf(525.74 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Estimating the result size of complex queries that involve selection on multiple attributes and the join of several relations is a difficult but fundamental task in database query processing. It arises in cost-based query optimization, query profiling, and approximate query answering. In this paper, we show how probabilistic graphical models can be effectively used for this task as an accurate and compact approximation of the joint frequency distribution of multiple attributes across multiple ...

29 Reports from KDD-2001: KDD Cup 2001 report

Jie Cheng, Christos Hatzis, Hisashi Hayashi, Mark-A. Krogel, Shinichi Morishita, David Page, Jun Sese

January 2002 **ACM SIGKDD Explorations Newsletter**, Volume 3 Issue 2Full text available:  pdf(1.96 MB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

This paper presents results and lessons from KDD Cup 2001. KDD Cup 2001 focused on mining biological databases. It involved three cutting-edge tasks related to drug design and genomics.

Keywords: Competition, biology, drug design, genomics

30 Evolving data mining into solutions for insights: Scaling mining algorithms to large databases

Paul Bradley, Johannes Gehrke, Raghu Ramakrishnan, Ramakrishnan Srikant

August 2002 **Communications of the ACM**, Volume 45 Issue 8Full text available:  pdf(116.66 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#) html(28.54 KB)

Which insights about data structure make it possible to analyze the very large databases collected by Internet, business, scientific, and government applications?

31 Special issue on ICML: Learning probabilistic models of link structure

Lisa Getoor, Nir Friedman, Daphne Koller, Benjamin Taskar

March 2003 **The Journal of Machine Learning Research**, Volume 3Full text available:  pdf(479.67 KB)Additional Information: [full citation](#), [abstract](#), [index terms](#)

Most real-world data is heterogeneous and richly interconnected. Examples include the Web, hypertext, bibliometric data and social networks. In contrast, most statistical learning methods work with "flat" data representations, forcing us to convert our data into a form that loses much of the link structure. The recently introduced framework of *probabilistic relational models* (PRMs) embraces the object-relational nature of structured data by capturing probabilistic interactions between att ...

32 SPARTAN: a model-based semantic compression system for massive data tables

Shivnath Babu, Minos Garofalakis, Rajeev Rastogi

May 2001 **ACM SIGMOD Record , Proceedings of the 2001 ACM SIGMOD international conference on Management of data**, Volume 30 Issue 2Full text available:  pdf(240.19 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

While a variety of lossy compression schemes have been developed for certain forms of digital data (e.g., images, audio, video), the area of lossy compression techniques for arbitrary data tables has been left relatively unexplored. Nevertheless, such techniques are clearly motivated by the ever-increasing data collection rates of modern enterprises and the need for effective, guaranteed-quality approximate answers to queries over massive relational data sets. In this paper, we propose SPA ...

33 Efficient discovery of error-tolerant frequent itemsets in high dimensions 

Cheng Yang, Usama Fayyad, Paul S. Bradley

August 2001 **Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  pdf(1.11 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

We present a generalization of frequent itemsets allowing for the notion of errors in the itemset definition. We motivate the problem and present an efficient algorithm that identifies error-tolerant frequent clusters of items in transactional data (customer-purchase data, web browsing data, text, etc.). The algorithm exploits sparseness of the underlying data to find large groups of items that are correlated over database records (rows). The notion of transaction coverage allows us to extend th ...

Keywords: Error-tolerant frequent itemset, clustering, collaborative filtering, high dimensions, query selectivity estimation

34 Probabilistic object bases 

Thomas Eiter, James J. Lu, Thomas Lukasiewicz, V. S. Subrahmanian

September 2001 **ACM Transactions on Database Systems (TODS)**, Volume 26 Issue 3

Full text available:  pdf(663.73 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Although there are many applications where an object-oriented data model is a good way of representing and querying data, current object database systems are unable to handle objects whose attributes are uncertain. In this article, we extend previous work by Kornatzky and Shimony to develop an algebra to handle object bases with uncertainty. We propose concepts of consistency for such object bases, together with an NP-completeness result, and classes of probabilistic object bases for which consi ...

Keywords: Consistency, object-oriented database, probabilistic object algebra, probabilistic object base, probability, query language, query optimization

35 Spatial decision support system for land assessment 

Cláudio Chauke Nehme, Margareth Simões

November 1999 **Proceedings of the seventh ACM international symposium on Advances in geographic information systems**

Full text available:  pdf(117.01 KB)

Additional Information: [full citation](#), [references](#), [index terms](#)

Keywords: GIS, agriculture planning, artificial intelligence, decision support system, expert system, geoinformatics, geoinformation system, land evaluation, land use planning

36 Research track: Probabilistic discovery of time series motifs 

Bill Chiu, Eamonn Keogh, Stefano Lonardi

August 2003 **Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  pdf(443.69 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Several important time series data mining problems reduce to the core task of finding approximately repeated subsequences in a longer time series. In an earlier work, we formalized the idea of approximately repeated subsequences by introducing the notion of time series motifs. Two limitations of this work were the poor scalability of the motif discovery algorithm, and the inability to discover motifs in the presence of noise. Here we address these limitations by introducing a novel algorithm insp ...

Keywords: data mining, motifs, randomized algorithms, time series

37 Accepted Posters: Information filtering using bayesian networks: effective user interfaces for aviation weather data 

Corinne Clinton Ruokangas, Ole J. Mengshoel

January 2003 **Proceedings of the 8th international conference on Intelligent user interfaces**

Full text available:  pdf(1.09 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Weather is a complex, dynamic process with tremendous impact on aviation. While pilots often have access to large amounts of aviation weather data, they find it difficult and time-consuming to identify weather hazards, due to the sheer amount and cryptic formatting of the data. To address this challenge, we have developed information filtering concepts based on a unified Bayesian network model, integrating text and graphical weather data in the context of specific mission, equipment and personal ...

Keywords: bayesian models, bayesian networks, data filtering, information management, intelligent visualization, situation awareness

38 Research sessions: query optimization: CORDS: automatic discovery of correlations and soft functional dependencies 

Ihab F. Ilyas, Volker Markl, Peter Haas, Paul Brown, Ashraf Aboulnaga

June 2004 **Proceedings of the 2004 ACM SIGMOD international conference on Management of data**

Full text available:  pdf(559.35 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

The rich dependency structure found in the columns of real-world relational databases can be exploited to great advantage, but can also cause query optimizers---which usually assume that columns are statistically independent---to underestimate the selectivities of conjunctive predicates by orders of magnitude. We introduce CORDS, an efficient and scalable tool for automatic discovery of correlations and soft functional dependencies between columns. CORDS searches for column pairs that might have ...

39 Compressed data cubes for OLAP aggregate query approximation on continuous dimensions 

Jayavel Shanmugasundaram, Usama Fayyad, P. S. Bradley

August 1999 **Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  pdf(1.12 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: OLAP, approximate query answering, clustering, data cubes, data mining, density estimation

40 Constraints in data mining: *SPARTAN*: using constrained models for guaranteed-error semantic compression

Shivnath Babu, Minos Garofalakis, Rajeev Rastogi
June 2002 **ACM SIGKDD Explorations Newsletter**, Volume 4 Issue 1

Full text available:  [pdf\(259.12 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

While a variety of lossy compression schemes have been developed for certain forms of digital data (e.g., images, audio, video), the area of lossy compression techniques for arbitrary data tables has been left relatively unexplored. Nevertheless, such techniques are clearly motivated by the ever-increasing data collection rates of modern enterprises and the need for effective, guaranteed-quality approximate answers to queries over massive relational data sets. In this paper, we propose *SPARTAN* ...

Results 21 - 40 of 200

Result page: [previous](#) [1](#) **2** [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2004 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)

USPTO PATENT FULL-TEXT AND IMAGE DATABASE[Home](#)[Quick](#)[Advanced](#)[Pat Num](#)[Help](#)[Bottom](#)[View Cart](#)

Searching 1976 to present...

Results of Search in 1976 to present db for:

(("Bayesian network" AND database) AND "expectation maximization"): 8 patents.

Hits 1 through 8 out of 8

[Jump To](#)[Refine Search](#)

"Bayesian network" and database and "expectation ma

PAT. NO. Title1 [6,701,016](#) **T** Method of learning deformation models to facilitate pattern matching2 [6,616,529](#) **T** Simulation and synthesis of sports matches3 [6,556,958](#) **T** Fast clustering with sparse data4 [6,529,891](#) **T** Automatic determination of the number of clusters by mixtures of bayesian networks5 [6,496,816](#) **T** Collaborative filtering with mixtures of bayesian networks6 [6,408,290](#) **T** Mixtures of bayesian networks with decision graphs7 [6,345,265](#) **T** Clustering with mixtures of bayesian networks8 [6,336,108](#) **T** Speech recognition with mixtures of bayesian networks[Top](#)[View Cart](#)[Home](#)[Quick](#)[Advanced](#)[Pat Num](#)[Help](#)